

D-vine copula based quantile regression

Daniel Kraus* and Claudia Czado†

April 20, 2016

Abstract

Quantile regression, that is the prediction of a random variable's quantiles conditioned on other random variables taking on certain values, has steadily gained importance in statistical modeling and financial applications. We introduce a new semiparametric quantile regression method based on sequentially fitting a likelihood optimal D-vine copula to given data resulting in highly flexible models with easily extractable conditional quantiles. As a subclass of regular vine copulas, D-vines enable the modeling of multivariate copulas in terms of bivariate building blocks, a so-called pair-copula construction (PCC). The proposed algorithm works fast and accurate even in high dimensions and incorporates an automatic variable selection by maximizing the conditional log-likelihood. Further, typical issues of quantile regression such as quantile crossing or transformations, interactions and collinearity of variables are automatically taken care of. In a simulation study the improved accuracy and saved computational time of the approach in comparison with established quantile regression methods is highlighted. An extensive financial application to international credit default swap (CDS) data including stress testing and Value-at-Risk (VaR) prediction demonstrates the usefulness of the proposed method.

Keywords: quantile regression, conditional distribution, vine copula, conditional copula quantile, stress testing

1 Introduction

Predicting quantiles (e.g. median or quartiles) of a random variable conditioned on other variables taking on fixed values, has continually attracted interest and found applications in various fields, especially in finance. It has become a standard tool for risk managers working on portfolio optimization, asset pricing and the evaluation of systemic risk. For example, [Adrian and Brunnermeier \(2011\)](#) introduce the *CoVaR*, a measure for systemic risk calculating conditional quantiles of a financial institution's loss distribution conditional on other institutions being in distress, and use it to evaluate the institution's contribution to systemic risk. A similar approach to measure systemic risk is found in [Brownlees and Engle \(2012\)](#). Further applications of quantile regression in the financial sector include measuring dependence in the FX markets

*Corresponding author, Zentrum Mathematik, Technische Universität München, Boltzmanstraße 3, 85748 Garching (email: daniel.kraus@tum.de)

†Zentrum Mathematik, Technische Universität München, Boltzmanstraße 3, 85748 Garching

(Bouyé and Salmon 2009), developing pricing models for real estates (Li et al. 2013) and predicting volatilities in the stock market (Noh et al. 2015).

The literature is quite rich in methods to predict conditional quantiles. The most famous and therefore frequently used method is linear quantile regression (Koenker and Bassett Jr 1978) which can be seen as the expansion of the well known ordinary least squares estimation used to predict conditional means. These simple linear models have been refined to account for nonparametric effects via additive models (Koenker et al. 2011, Fenske et al. 2012). Further methods include local quantile regression (Spokoiny et al. 2013), single-index quantile regression (Wu et al. 2010), semiparametric quantile regression (Noh et al. 2015), nonparametric quantile regression (Li et al. 2013) and quantile regression for time series (e.g. Chen et al. 2009, Xiao and Koenker 2012). In the machine learning context, Hwang and Shim (2005) use support vector machines for conditional quantile estimation while random forests are utilized in Meinshausen (2006). Moreover, Bouyé and Salmon (2009) propose a general approach to nonlinear quantile regression with one predictor based on a copula function.

In this paper we go beyond the existing methodologies and use the flexible class of D-vine copula models for the prediction of conditional quantiles. D-vines are a subclass of regular vine copulas which have enhanced statistical modeling since the publication of the seminal paper of Aas et al. (2009). Vines enable the modeling of multivariate copula densities in terms of bivariate building blocks resulting in versatile parametric models. Given a response and predictor variables, we fit a D-vine to the data that maximizes the conditional likelihood of the response given the predictors by sequentially adding covariates to the model until the remaining covariates are deemed as non-influential. Thus, the question of variable selection is answered automatically. As we will show, the use of D-vine copulas allows us to extract the desired conditional quantiles precisely and without computational effort in a way that naturally avoids quantile crossing.

The remainder of the paper is organized as follows. Section 2 introduces the concept of D-vines while Section 3 describes how they can be used for the prediction of conditional quantiles. We demonstrate the usefulness of the approach in a simulation study. The prediction performance of D-vine quantile regression and its competitor methods, which are discussed in Section 4, is presented in Section 5. For the demonstration of the usefulness of our method we include a real data example in Section 6 that contains financial applications of the D-vine quantile regression featuring stress testing and Value-at-Risk prediction. Section 7 gives conclusions and areas of future research.

2 D-vine copulas

This section derives the main tools needed for D-vine based quantile regression. First, we introduce the notion of a *copula*, a probabilistic device that has evolved to be the standard instrument for dependence modeling over the last years. A d -dimensional copula C is a d -variate distribution function on the unit hypercube $[0, 1]^d$ with uniform marginal distribution functions. Sklar’s Theorem (Sklar 1959) provides a link between multivariate distributions and their associated copulas. It states that for every multivariate random vector $\mathbf{X} = (X_1, \dots, X_d)' \sim F$ with marginal distribution functions F_1, \dots, F_d , there exists a copula C associated with \mathbf{X} , such that

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)).$$

This decomposition of the multivariate distribution into its margins and its associated copula is unique when \mathbf{X} is absolutely continuous (which we will assume for the

remainder of this paper). In that case, the density of \mathbf{X} can be decomposed similarly:

$$f(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d))f_1(x_1) \cdot \dots \cdot f_d(x_d),$$

where $c(u_1, \dots, u_d) := \frac{\partial^d}{\partial_1 \dots \partial_d} C(u_1, \dots, u_d)$ is the copula density and f_1, \dots, f_d are the marginal densities.

Often, when one's interest is focused only on the dependence structure of a random vector $(X_1, \dots, X_d)'$ disregarding its marginal distributions, the random variables are considered on the so-called *u-scale* (or copula scale) by applying the *probability integral transform* (PIT) to them:

$$U_j := F_j(X_j), \quad j = 1, \dots, d. \quad (2.1)$$

The U_j are then uniformly distributed and their joint distribution function is the copula C associated with \mathbf{X} . This implies that the copulas associated with \mathbf{X} and \mathbf{U} coincide. Refer to [Joe \(1997\)](#) and [Nelsen \(2007\)](#) for a more thorough examination of copulas including a number of examples of parametric copulas, especially bivariate *pair-copulas*. Those are of special interest to us since they are the building blocks used for the pair-copula construction of D-vines.

For a random vector $(X_1, \dots, X_d)'$, a set $D \subset \{1, \dots, d\}$ and $i, j \in \{1, \dots, d\} \setminus D$ we use the following notations:

- (a) $C_{X_i, X_j; \mathbf{x}_D}(\cdot, \cdot; \mathbf{x}_D)$ denotes the copula associated with the conditional distribution of $(X_i, X_j)'$ given $\mathbf{X}_D = \mathbf{x}_D$. We abbreviate this by $C_{ij;D}(\cdot, \cdot; \mathbf{x}_D)$. Further, $c_{ij;D}(\cdot, \cdot; \mathbf{x}_D)$ is the copula density corresponding to $C_{ij;D}(\cdot, \cdot; \mathbf{x}_D)$.
- (b) $F_{X_i|\mathbf{x}_D}(\cdot, \cdot; \mathbf{x}_D)$ denotes the conditional distribution of the random variable X_i given $\mathbf{X}_D = \mathbf{x}_D$. We use $F_{i|D}(\cdot, \cdot; \mathbf{x}_D)$ as an abbreviation.
- (c) $C_{U_i|\mathbf{U}_D}(\cdot, \cdot; \mathbf{u}_D)$ denotes the conditional distribution of the PIT random variable U_i given $\mathbf{U}_D = \mathbf{u}_D$. We abbreviate this by $C_{i|D}(\cdot, \cdot; \mathbf{u}_D)$.

Following [Czado \(2010\)](#), the joint density f of the continuously distributed random vector \mathbf{X} can be written in terms of (conditional) bivariate copula densities and its marginal densities as

$$f(x_1, \dots, x_d) = \prod_{k=1}^d f_k(x_k) \prod_{i=1}^{d-1} \prod_{j=i+1}^d c_{ij|i+1, \dots, j-1}(F_{i|i+1, \dots, j-1}(x_i|x_{i+1}, \dots, x_{j-1}), F_{j|i+1, \dots, j-1}(x_j|x_{i+1}, \dots, x_{j-1}); x_{i+1}, \dots, x_{j-1}). \quad (2.2)$$

We call this pair-copula construction (PCC) a *D-vine density* with order $1-2-\dots-d$. If all margins are uniform, we speak of a *D-vine copula*. As introduced by [Bedford and Cooke \(2002\)](#), we present a graph theoretic representation of the D-vine. This helps to visualize the dependencies and organize the pair-copulas appearing in [Equation \(2.2\)](#), since each edge of the graph corresponds to a certain pair-copula.

Example 2.1. [Figure 1](#) shows an exemplary 5-dimensional D-vine corresponding with

$$\begin{aligned} f(x_1, x_2, x_3, x_4, x_5) &= f_1(x_1)f_2(x_2)f_3(x_3)f_4(x_4)f_5(x_5) \\ &\quad \cdot c_{12} \cdot c_{23} \cdot c_{34} \cdot c_{45} & (T_1) \\ &\quad \cdot c_{13;2} \cdot c_{24;3} \cdot c_{35;4} & (T_2) \\ &\quad \cdot c_{14;23} \cdot c_{25;34} & (T_3) \\ &\quad \cdot c_{15;234}, & (T_4) \end{aligned}$$

where for brevity we omitted the arguments of the pair-copulas.

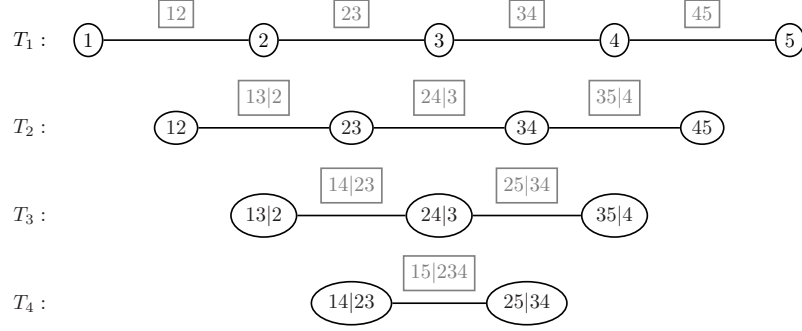


Figure 1: Graph theoretic representation of a D-vine with order 1–2–3–4–5. The nodes of the trees are plotted in black circles and the corresponding pair-copulas in gray squares.

We see that all pair-copulas used in the decomposition appear as edges in the corresponding nested set of trees displayed in Figure 1.

In order to fit a D-vine copula with a fixed order to given data, all pair-copulas appearing in Equation (2.2) are estimated as parametric bivariate copulas (see Algorithm A.2). A common assumption when working with D-vines is to assume that the copulas associated with conditional distributions $c_{i,j|D}$ do not depend on the specific values of the conditioning vector \mathbf{x}_D , i.e. $c_{i,j|D}(\cdot, \cdot; \mathbf{x}_D) \equiv c_{i,j|D}(\cdot, \cdot)$. Nevertheless, the arguments of the copulas, $F_{i|D}(x_i|\mathbf{x}_D)$ and $F_{j|D}(x_j|\mathbf{x}_D)$, still do depend on the conditioning vector. A more detailed examination of this so-called simplifying assumption can for example be found in Stöber et al. (2013), Haff et al. (2010) and Killiches et al. (2016).

The conditional distributions of the form $F_{i|D}(x_i|\mathbf{x}_D)$ appearing in the PCC can be evaluated using only the pair-copulas specified for the D-vine from lower trees by applying the following recursion, which was first stated by Joe (1997): Let $l \in D$ and $D_{-l} := D \setminus \{l\}$. Then,

$$F_{i|D}(x_i|\mathbf{x}_D) = h_{i|l;D_{-l}}(F_{i|D_{-l}}(x_i|\mathbf{x}_{D_{-l}}) | F_{l|D_{-l}}(x_l|\mathbf{x}_{D_{-l}})), \quad (2.3)$$

where for $i, j \notin D, i < j$, $h_{i|j;D}(u|v) := \partial C_{ij;D}(u, v) / \partial v = C_{i|j;D}(u|v)$ and $h_{j|i;D}(u|v) := \partial C_{ij;D}(u, v) / \partial u = C_{j|i;D}(u|v)$ are the h-functions associated with the pair-copula $C_{ij;D}$. Therefore, continuing our 5-dimensional example, e.g. the first argument of the copula $c_{14;23}$ from Tree 3, namely $F_{1|23}(x_1|x_2, x_3)$, can be evaluated using the h-function associated with $C_{13;2}$ from Tree 2 and the h-functions associated with C_{12} and C_{23} from Tree 1, such that

$$\begin{aligned} F_{1|23}(x_1|x_2, x_3) &= h_{1|3;2}(F_{1|2}(x_1|x_2) | F_{3|2}(x_3|x_2)) \\ &= h_{1|3;2}(h_{1|2}(F_1(x_1) | F_2(x_2)) | h_{3|2}(F_3(x_3) | F_2(x_2))). \end{aligned}$$

Note that there were two possibilities for the choice of l in Equation (2.3) when simplifying $F_{1|23}$, $l = 2$ and $l = 3$. We chose $l = 3$ and ended up with the h-function associated with $C_{13;2}$. Had we chosen $l = 2$, we would have needed the h-function associated with $C_{12;3}$, which is not a copula specified by the D-vine (cf. Figure 1). In general, with the appropriate choice of l in each step we can express all appearing conditional distributions in terms of h-functions associated with the pair-copulas of the D-vine.

3 D-vine based quantile regression model

3.1 Conditional quantile function

The main purpose of D-vine copula based quantile regression is to predict the quantile of a response variable Y given the outcome of some predictor variables X_1, \dots, X_d , $d \geq 1$, where $Y \sim F_Y$ and $X_j \sim F_j$, $j = 1, \dots, d$. Hence, the focus of interest lies on the joint modeling of Y and \mathbf{X} and in particular on the conditional quantile function for $\alpha \in (0, 1)$:

$$q_\alpha(x_1, \dots, x_d) := F_{Y|X_1, \dots, X_d}^{-1}(\alpha|x_1, \dots, x_d). \quad (3.4)$$

Using the probability integral transforms $V := F_Y(Y)$ and $U_j := F_j(X_j)$ with corresponding PIT values $v := F_Y(y)$ and $u_j := F_j(x_j)$, it follows that

$$\begin{aligned} F_{Y|X_1, \dots, X_d}(y|x_1, \dots, x_d) &= P(Y \leq y|X_1 = x_1, \dots, X_d = x_d) \\ &= P(F_Y(Y) \leq v|F_1(X_1) = u_1, \dots, F_d(X_d) = u_d) \\ &= C_{V|U_1, \dots, U_d}(v|u_1, \dots, u_d). \end{aligned}$$

Therefore, inversion yields

$$F_{Y|X_1, \dots, X_d}^{-1}(\alpha|x_1, \dots, x_d) = F_Y^{-1}\left(C_{V|U_1, \dots, U_d}^{-1}(\alpha|u_1, \dots, u_d)\right). \quad (3.5)$$

Hence, the conditional quantile function can be expressed in terms of the inverse marginal distribution function of Y and the *conditional copula quantile function* $C_{V|U_1, \dots, U_d}^{-1}$ conditioned on the PIT values of \mathbf{x} . Note that this result was already stated for the one-dimensional predictor case in Equation (2) of [Bernard and Czado \(2015\)](#). Now, we can obtain an estimate of the conditional quantile function by estimating the marginals F_Y and F_j , $j = 1, \dots, d$, as well as the copula C_{V, U_1, \dots, U_d} and plugging them into [Equation \(3.5\)](#):

$$\hat{q}_\alpha(x_1, \dots, x_d) := \hat{F}_Y^{-1}\left(\hat{C}_{V|U_1, \dots, U_d}^{-1}(\alpha|\hat{u}_1, \dots, \hat{u}_d)\right), \quad (3.6)$$

where $\hat{u}_j := \hat{F}_j(x_j)$ is the estimated PIT of x_j , $j = 1, \dots, d$.

While regarding the \hat{F}_j there is a vast literature about the estimation of a univariate CDF, the question arises how to estimate the multivariate copula C_{V, U_1, \dots, U_d} , such that on the one hand it facilitates a flexible model that is able to capture asymmetric dependencies, heavy tails and tail dependencies between the variables, and on the other hand the estimated conditional quantile function $\hat{C}_{V|U_1, \dots, U_d}^{-1}(\alpha|\hat{u}_1, \dots, \hat{u}_d)$ is easily calculable. The answer we suggest is to fit a D-vine copula to $(V, U_1, \dots, U_d)'$, such that V is the first node in the first tree (i.e. a D-vine with order $V-U_{l_1}-\dots-U_{l_d}$, where $(l_1, \dots, l_d)'$ is allowed to be an arbitrary permutation of $(1, \dots, d)'$). This results in a flexible class of copulas since each bivariate copula of the pair-copula construction can be modeled separately and the order of the U_j can be chosen such that the conditional likelihood is maximized as will be explained in detail in the next section. Finally, the recursion given in [Equation \(2.3\)](#) allows us to express $C_{V|U_1, \dots, U_d}(v|u_1, \dots, u_d)$ in terms of nested h-functions and consequently, $C_{V|U_1, \dots, U_d}^{-1}(\alpha|u_1, \dots, u_d)$ in terms of inverse h-functions. We will now demonstrate this in a 4-dimensional example.

Example 3.1. For a D-vine with order $V-U_1-U_2-U_3$, the conditional distribution of

V given $(U_1, U_2, U_3)'$ can recursively be expressed as

$$\begin{aligned} C_{V|U_1, U_2, U_3}(v|u_1, u_2, u_3) &= h_{V|U_3; U_1, U_2}(C_{V|U_1, U_2}(v|u_1, u_2)|C_{U_3|U_1, U_2}(u_3|u_1, u_2)) \\ &= h_{V|U_3; U_1, U_2}(h_{V|U_2; U_1}(C_{V|U_1}(v|u_1)|C_{U_2|U_1}(u_2|u_1))| \\ &\quad h_{U_3|U_1; U_2}(C_{U_3|U_2}(u_3|u_2)|C_{U_1|U_2}(u_1|u_2))) \\ &= h_{V|U_3; U_1, U_2}(h_{V|U_2; U_1}(h_{V|U_1}(v|u_1)|h_{U_2|U_1}(u_2|u_1))| \\ &\quad h_{U_3|U_1; U_2}(h_{U_3|U_2}(u_3|u_2)|h_{U_1|U_2}(u_1|u_2))) \end{aligned}$$

Inversion yields the conditional quantile function:

$$\begin{aligned} C_{V|U_1, U_2, U_3}^{-1}(\alpha|u_1, u_2, u_3) &= \\ h_{V|U_1}^{-1} \left[h_{V|U_2; U_1}^{-1} \left\{ h_{V|U_3; U_1, U_2}^{-1} \left(\alpha | h_{U_3|U_1; U_2}(h_{U_3|U_2}(u_3|u_2)|h_{U_1|U_2}(u_1|u_2)) \right) | h_{U_2|U_1}(u_2|u_1) \right\} | u_1 \right] \end{aligned}$$

Note that $C_{V|U_1, \dots, U_d}^{-1}(\alpha|u_1, \dots, u_d)$ is monotonically increasing in α . Therefore, a crossing of quantile functions corresponding to different quantile levels is not possible. This issue of quantile crossing often arises in linear and non-linear quantile regression (e.g. see the application section of [Fenske et al. \(2012\)](#)). [Bernard and Czado \(2015\)](#) show that in linear regression quantile functions cross if non-Gaussian data is modeled. In addition, in (non-)linear quantile regression a substantial amount of effort has to be put into dealing with issues such as transformation of response and covariates, including interactions among covariates and collinearity between covariates. Our approach solves these issues automatically since the distribution class given by the D-vines is much more flexible and makes less restrictive model assumptions how the covariates influence the response. This is also noted for regular vine regression by [Cooke et al. \(2015\)](#).

3.2 Estimation process

This section explains how the estimate of $q_\alpha(\mathbf{x})$ is obtained using a two step estimation procedure. All numerical calculations are done using the programming language R ([R Core Team 2015](#)) using the package `VineCopula` ([Schepsmeier et al. 2015](#)).

Let $\mathbf{y} := (y^{(i)})_{i=1, \dots, n}$, $\mathcal{X} := (x_j^{(i)})_{j=1, \dots, d, i=1, \dots, n}$ be n independent and identically distributed observations of the random vector $(Y, X_1, X_2, \dots, X_d)'$. The representation of $\hat{q}_\alpha(\mathbf{x})$ in [Equation \(3.6\)](#) allows us to divide the estimation process into two steps. In the first step we estimate the marginal distribution functions F_Y and F_j of Y and X_j , $j = 1, \dots, d$, respectively, and in the second step the D-vine that specifies the pair copulas needed to evaluate $\hat{C}_{V|U_1, \dots, U_d}^{-1}(\alpha|\hat{u}_1, \dots, \hat{u}_d)$ is estimated.

First step: estimation of the marginals

In general, we have two choices of how to fit the marginal distributions, either parametrically or nonparametrically. Since we will fit the copula in the second step parametrically, this choice will either result in a fully parametric or semiparametric estimate of $q_\alpha(\mathbf{x})$. [Noh et al. \(2013\)](#) point out that modeling the marginals as well as the copula parametrically might cause the resulting fully parametric estimator to be biased and inconsistent if one of the parametric models is misspecified. Therefore we prefer the semiparametric approach and estimate the marginals by the kernel smoothing estimator ([Parzen 1962](#)), which is, given a sample $(x^{(i)})_{i=1, \dots, n}$, defined as

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - x^{(i)}}{h}\right), \quad x \in \mathbb{R}, \quad (3.7)$$

where $K(x) := \int_{-\infty}^x k(t)dt$ with $k(\cdot)$ being a symmetric probability density function and $h > 0$ a bandwidth parameter. Usually, we choose $k = \varphi$, i.e. a Gaussian kernel, and the plug-in bandwidth developed in [Duong \(2015b\)](#), Equation (4), which minimizes the asymptotic mean integrated squared error (implemented in the function `kcde` of the package `ks` ([Duong 2015a](#))).

Hence, we obtain \hat{F}_Y and \hat{F}_j as estimates for the marginal distribution functions. We use these to transform the observed data to pseudo copula data $\hat{v}^{(i)} := \hat{F}_Y(y^{(i)})$ and $\hat{u}_j^{(i)} := \hat{F}_j(x_j^{(i)})$, $j = 1, \dots, d$, $i = 1, \dots, n$. The pseudo copula data $\hat{\mathbf{v}} = (\hat{v}^{(i)})_{i=1, \dots, n}$, $\hat{\mathcal{U}} = (\hat{u}_j^{(i)})_{j=1, \dots, d, i=1, \dots, n}$ is then an approximately i.i.d. sample from the PIT random vector $(V, U_1, \dots, U_d)'$ and will be used to estimate the D-vine copula in the second step.

Second step: estimation of the D-vine

As motivated by [Equation \(3.6\)](#), we fit a D-vine with order $V-U_{l_1}-\dots-U_{l_d}$ to the pseudo copula data, since then the evaluation of $\hat{C}_{V|U_1, \dots, U_d}^{-1}(\alpha|\hat{u}_1, \dots, \hat{u}_d)$ which is needed to calculate the conditional quantile is easily feasible. The ordering $\mathbf{l} = (l_1, \dots, l_d)'$ can generally be chosen arbitrary. However, we want to choose it such that the resulting model for the prediction of the conditional quantile has a high explanatory power. Since it would be infeasible to compare all $d!$ possible orderings, we construct the D-vine sequentially. Starting with zero covariates, in each step we add the covariate to the model that improves the model's fit the most. As a measure for the model's fit, we define the conditional log-likelihood (cll) of an estimated D-vine copula with order \mathbf{l} , estimated parametric pair-copula families $\hat{\mathcal{F}}$ and corresponding copula parameters $\hat{\boldsymbol{\theta}}$ given pseudo copula data $(\hat{\mathbf{v}}, \hat{\mathcal{U}})$ as

$$\text{cll}(\mathbf{l}, \hat{\mathcal{F}}, \hat{\boldsymbol{\theta}}; \hat{\mathbf{v}}, \hat{\mathcal{U}}) := \sum_{i=1}^n \ln c_{V|\mathbf{U}}(\hat{v}^{(i)}|\hat{\mathbf{u}}^{(i)}; \mathbf{l}, \hat{\mathcal{F}}, \hat{\boldsymbol{\theta}}). \quad (3.8)$$

The conditional copula density $\hat{c}_{V|\mathbf{U}}$ can be expressed as the product over all pair-copulas of the D-vine that contain V (see [Killiches et al. 2015](#)):

$$\begin{aligned} c_{V|\mathbf{U}}(\hat{v}^{(i)}|\hat{\mathbf{u}}^{(i)}; \mathbf{l}, \hat{\mathcal{F}}, \hat{\boldsymbol{\theta}}) &= c_{VU_{l_1}}(v^{(i)}, u_{l_1}^{(i)}; \hat{\mathcal{F}}_{VU_{l_1}}, \hat{\boldsymbol{\theta}}_{VU_{l_1}}) \times \\ &\prod_{j=2}^n c_{VU_{l_j}; U_{l_1}, \dots, U_{l_{j-1}}}(\hat{C}_{V|U_{l_1}, \dots, U_{l_{j-1}}}(\hat{v}^{(i)}|u_{l_1}^{(i)}, \dots, u_{l_{j-1}}^{(i)}), \hat{C}_{U_{l_j}|U_{l_1}, \dots, U_{l_{j-1}}}(\hat{u}_{l_1}^{(i)}|u_{l_1}^{(i)}, \dots, u_{l_{j-1}}^{(i)}); \\ &\quad \hat{\mathcal{F}}_{VU_{l_j}; U_{l_1}, \dots, U_{l_{j-1}}}, \hat{\boldsymbol{\theta}}_{VU_{l_j}; U_{l_1}, \dots, U_{l_{j-1}}}), \end{aligned}$$

where $\hat{\mathcal{F}}_I$ and $\hat{\boldsymbol{\theta}}_I$ denote the family and parameter of pair-copula \hat{c}_I .

We now describe the D-vine regression algorithm which sequentially constructs a D-vine while maximizing the model's conditional log-likelihood in each step (for a detailed code see [Algorithm A.1](#) in [Appendix A](#)). Assume that at the beginning of the k th step of the algorithm the current optimal D-vine contains $k-1$ predictors (for illustration, see the black D-vine in [Figure 2](#)). For each of the remaining variables U_j that have not been chosen yet, we fit the pair-copulas that are needed to extend the current model to a D-vine with order $V-U_{l_1}-\dots-U_{l_{k-1}}-U_j$ (see the gray circles) and compute the resulting model's conditional log-likelihood. Then, the current model is updated by adding the variable corresponding to the highest cll, concluding step k . That way, step by step the covariates are ordered regarding their power to predict the response.

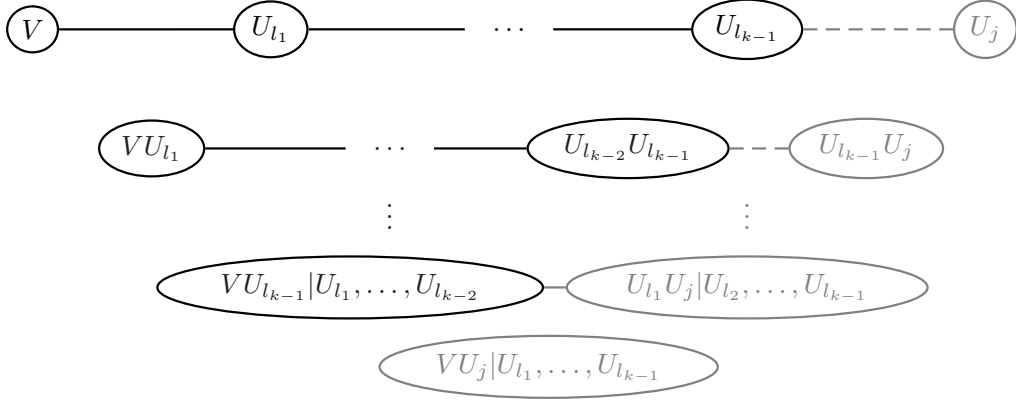


Figure 2: Extending the current D-vine (black) by adding U_j in the k -th step of the algorithm. For this purpose, the gray pair-copulas have to be estimated.

In the case that in the k th step none of the remaining covariates is able to increase the model's cll, the algorithm stops and returns the model only containing the $k - 1$ chosen covariates so far. Therefore, an *automatic forward covariate selection* is accomplished resulting in parsimonious models. In order to get even more parsimonious models, we also consider two variants of the cll penalizing the number of parameters $|\hat{\theta}|$ used for the construction of the D-vine: the AIC-corrected conditional log-likelihood cll^{AIC} , defined as

$$\text{cll}^{\text{AIC}}(\mathbf{l}, \hat{\mathcal{F}}, \hat{\theta}; \hat{\mathbf{v}}, \hat{\mathcal{U}}) := -2 \text{cll}(\mathbf{l}, \hat{\mathcal{F}}, \hat{\theta}; \hat{\mathbf{v}}, \hat{\mathcal{U}}) + 2|\hat{\theta}|$$

and the BIC-corrected conditional log-likelihood cll^{BIC} , defined as

$$\text{cll}^{\text{BIC}}(\mathbf{l}, \hat{\mathcal{F}}, \hat{\theta}; \hat{\mathbf{v}}, \hat{\mathcal{U}}) := -2 \text{cll}(\mathbf{l}, \hat{\mathcal{F}}, \hat{\theta}; \hat{\mathbf{v}}, \hat{\mathcal{U}}) + \log(n)|\hat{\theta}|.$$

Depending on how parsimonious the resulting model is desired to be, one can decide which version of the conditional log-likelihood to use. In our applications in the later sections we always use the AIC-corrected cll^{AIC} since in the simulation study it has shown to select the most reasonable models, i.e. disregarding unimportant variables while keeping the influential ones.

Example 3.2. We illustrate how the algorithm works for a four-dimensional data set $(y^{(i)}, x_1^{(i)}, x_2^{(i)}, x_3^{(i)})'$, $i = 1, \dots, n = 500$, sampled from $(Y, X_1, X_2, X_3)' \sim \mathcal{N}_4(\mathbf{0}, \Sigma)$ with

$$\Sigma = \begin{pmatrix} 1 & 0.4 & 0.8 & 0 \\ 0.4 & 1 & 0.32 & 0 \\ 0.8 & 0.32 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

First, the data is transformed to pseudo copula data $(\hat{v}^{(i)}, \hat{u}_1^{(i)}, \hat{u}_2^{(i)}, \hat{u}_3^{(i)})'$, $i = 1, \dots, n$, using the kernel smoothing estimators introduced in Equation (3.7).

In the **first step** of the algorithm, for each of the pairs $(V, U_j)'$, $j = 1, 2, 3$, the AIC-optimal pair-copula is chosen using the function `BiCopSelect` of the R package `VineCopula` with an independence test at level 0.05 (see Appendix A, Algorithm A.2). Further, the conditional log-likelihood is calculated for each of the pairs (we omit the AIC- and BIC-corrected cll-values here, since the fitted models have the same number of parameters and therefore these statistics would imply the same conclusions). The

Pair-copula	\hat{C}_{V,U_1}	\hat{C}_{V,U_2}	\hat{C}_{V,U_3}
Family	Gauss	Gauss	Indep
Parameter	0.34	0.79	0
cll	33.0	249.4	0

Table 1: Candidate models with corresponding cll after the algorithm’s first step.

results are shown in the [Table 1](#). Implying the largest cll, U_2 is chosen as the first variable to construct the D-vine.

In the **second step**, we investigate whether the addition of either of the remaining variables U_1 or U_3 to the current D-vine can improve the conditional log-likelihood of the model. Adding U_1 would update the D-vine to order $V-U_2-U_1$ with newly estimated pair-copulas $\hat{C}_{V,U_1;U_2}$ (Gaussian with $\rho = 0.27$) and \hat{C}_{U_1,U_2} (Gaussian with $\rho = 0.23$). The log-likelihood of the resulting conditional copula $\hat{c}_{V|U_2,U_1}$ is 269.8. The addition of U_3 would result in both new pair-copulas to be estimated as independence copulas. Consequently, the conditional log-likelihood would not improve compared to the model without U_3 . Since $269.8 > 249.4$, we update the vine to order $V-U_2-U_1$.

In the **third step**, we check whether the addition of the remaining variable U_3 to the D-vine improves the conditional log-likelihood of the model. Not surprisingly, as in the second step the new pair copulas $\hat{C}_{V,U_3;U_2,U_1}$, $\hat{C}_{U_2,U_3;U_1}$ and \hat{C}_{U_1,U_3} are all estimated to be the independence copula. Hence, the conditional log-likelihood of the full model with D-vine with order $V-U_2-U_1-U_3$ is equal to the one with order $V-U_2-U_1$. Consequently, the algorithm stops and returns the D-vine with order $V-U_2-U_1$.

Thus, we have seen that the algorithm automatically selects influential covariates ranking them by importance and disregards superfluous variables.

In the following simulation study, we compare the performance of D-vine based quantile regression to conventional and established quantile regression methods in settings where the true quantile is known.

4 Established quantile regression methods

Linear quantile regression (LQR) [Koenker and Bassett Jr \(1978\)](#) were one of the first researchers in the field of quantile regression. For their linear quantile regression model, they assume the predicted conditional quantiles to be *linear* in the predictors, i.e.

$$\hat{q}_\alpha \left(x_1^{(i)}, \dots, x_d^{(i)} \right) := \beta_0 + \sum_{j=1}^d \beta_j x_j^{(i)}, \quad (4.9)$$

where the regression coefficients $\beta \in \mathbb{R}^{d+1}$ are chosen to solve the minimization problem

$$\min_{\beta \in \mathbb{R}^{d+1}} \left\{ \alpha \sum_{i=1}^n \left(y^{(i)} - \beta_0 - \sum_{j=1}^d \beta_j x_j^{(i)} \right)^+ + (1 - \alpha) \sum_{i=1}^n \left(\beta_0 + \sum_{j=1}^d \beta_j x_j^{(i)} - y^{(i)} \right)^+ \right\}. \quad (4.10)$$

Linear quantile regression has drawn criticism for its strong assumption of linear conditional quantiles. [Bernard and Czado \(2015\)](#) highlight its pitfalls showing that the linearity assumption is almost never fulfilled. For example, assuming normal margins the only copula resulting in linear conditional quantiles is the Gaussian copula, which is very restrictive. Another major flaw of linear quantile regression they point out is

that regression lines of different quantile levels may cross, since they may have differing slopes. These drawbacks are also seen in the results of the simulation. Regarding implementation, we use the R-function `rq` of the package `quantreg` (Koenker 2013).

Boosting additive quantile regression (BAQR) In order to relax the linearity assumption of the above method Koenker (2005) proposes to use additive models for quantile regression, i.e.

$$\hat{q}_\alpha \left(x_1^{(i)}, \dots, x_d^{(i)}, z_1^{(i)}, \dots, z_J^{(i)} \right) := \beta_0 + \sum_{j=1}^d \beta_j x_j^{(i)} + \sum_{j=1}^J g_j \left(z_j^{(i)} \right), \quad (4.11)$$

where g_j denotes a smooth function of the continuous variable $z_j^{(i)}$, $j = 1, \dots, J$. Commonly, these functions are estimated by means of B-spline basis functions with a rather small number of knots at fixed points, solving a minimization similar to Equation (4.10) with additional penalization of the total variation of the derivatives of the smooth functions (see Koenker et al. 2011). In contrast, Fenske et al. (2012) use boosting techniques for the estimation of additive quantile regression and show in simulation studies the superiority over the total variation regularization approach. In short, boosting minimizes a given loss function (analogue to Equation (4.10)) by stepwise updating the estimator along the steepest gradient descent of the loss function. The optimal number of boosting iterations is determined as in Fenske et al. (2012) using a test data set. The algorithm is implemented in the function `gamboost` of the package `mboost` (Hothorn et al. 2015), for which we choose the parameter settings exactly as in the electronic supplement of Fenske et al. (2012).

Nonparametric quantile regression (NPQR) A fully nonparametric estimation of a conditional quantile is proposed by Li et al. (2013). After estimating the conditional distribution function $\hat{F}(y|x_1, \dots, x_d)$ nonparametrically applying a kernel estimator with an automatic data-driven bandwidth selector, they estimate the conditional quantile function $q_\alpha(x_1, \dots, x_d)$ by numerically inverting the estimated conditional distribution function, solving

$$\hat{q}_\alpha(x_1, \dots, x_d) := \underset{q}{\operatorname{argmin}} \left| \alpha - \hat{F}(q|x_1, \dots, x_d) \right|. \quad (4.12)$$

In R, we use the function `npqreg` of package `np` (Hayfield and Racine 2008), where the bandwidths are selected automatically by the function `npcdistbw` (with default settings).

Semiparametric quantile regression (SPQR) The semiparametric method of Noh et al. (2015) uses a copula density weighted quantile regression approach. Similar to our procedure, they propose a copula-based multivariate model for which they estimate the marginal distributions nonparametrically and the copula as a regular vine \hat{c} (see Aas et al. (2009)), which is a generalization of D-vines. Since in this case the conditional copula quantile function $C_{V|U_1, \dots, U_d}^{-1}$ in Equation (3.6) can in general not be calculated as elegantly as in the case of a fitted D-vine, they suggest to estimate the conditional quantile via minimizing a weighted check function:

$$\hat{q}_\alpha(x_1, \dots, x_d) = \underset{a}{\operatorname{argmin}} \sum_{i=1}^n \rho_\alpha(y^{(i)} - a) \hat{c} \left(\hat{F}_Y(y^{(i)}), \hat{F}_1(x_1^{(i)}), \dots, \hat{F}_d(x_d^{(i)}) \right), \quad (4.13)$$

where $\rho_\alpha(y) = y(\alpha - \mathbb{1}(y < 0))$ denotes the check function. Since this method is currently not implemented in R, we use the function `RVineStructureSelect` of the package `VineCopula` for the estimation of the regular vine and solve for the conditional quantile minimizing Equation (4.13) using the `optimize` function.

5 Results of the simulation study

We consider the following three scenarios for the random vector $(Y, X_1, \dots, X_d)'$:

- **C3**: $(Y, X_1, X_2)'$ follows a three-dimensional Clayton copula with parameter δ_1 or δ_2 and margin set \mathcal{M}_1 or \mathcal{M}_2 (see Table 2, row 1).
- **t5**: $(Y, X_1, \dots, X_4)'$ follows a five-dimensional t copula with 3 degrees of freedom, association matrix R_1 or R_2 and margin set \mathcal{M}_1 or \mathcal{M}_2 (see Table 2, row 2).
- **M5**: $\mathbf{X} \sim N_4(\mathbf{0}, \Sigma)$, with $\Sigma_{ij} = 0.5^{|i-j|}$.
 $Y := \sqrt{|2x_1 - x_2 + 0.5| + (-0.5x_3 + 1)(0.1x_4^3)} + \sigma\epsilon$ with $\epsilon \sim N(0, 1)$, $\sigma \in \{0.1, 1\}$.

Scen.	Copula parameter	Marginals			
C3	$\delta_1 = 0.86$ $\delta_2 = 4.67$		Y	X_1	X_2
		\mathcal{M}_1	$\mathcal{N}(0, 1)$	$t_4(0, 1)$	$\mathcal{N}(1, 4)$
		\mathcal{M}_2	$st_4(0, 1, 2)$	$s\mathcal{N}(-2, 0.5, 3)$	$st_3(1, 2, 5)$
t5	$R_1 = \begin{pmatrix} 1 & 0.6 & 0.5 & 0.5 & 0.4 \\ 0.6 & 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 & 0.5 \\ 0.4 & 0.5 & 0.5 & 0.5 & 1 \end{pmatrix}$		Y	X_1	X_2
		\mathcal{M}_1	$\mathcal{N}(0, 1)$	$t_4(0, 1)$	$\mathcal{N}(1, 4)$
		\mathcal{M}_2	$st_4(0, 1, 2)$	$s\mathcal{N}(-2, 0.5, 3)$	$st_3(1, 2, 5)$
	$R_2 = \begin{pmatrix} 1 & 0.27 & 0.74 & 0.72 & 0.41 \\ 0.27 & 1 & 0.28 & 0.29 & 0.27 \\ 0.74 & 0.28 & 1 & 0.74 & 0.42 \\ 0.72 & 0.29 & 0.74 & 1 & 0.40 \\ 0.41 & 0.27 & 0.42 & 0.40 & 1 \end{pmatrix}$		X_3	X_4	
		\mathcal{M}_1	$t_4(0, 1)$	$\mathcal{N}(1, 4)$	
		\mathcal{M}_2	$s\mathcal{N}(-2, 0.5, 3)$	$st_3(1, 2, 5)$	

Table 2: Parameter and marginal settings for Scenarios **C3** and **t5** of the simulation study. $(s\mathcal{N}(\mu, \sigma^2, \xi))$ and $st_\nu(\mu, \sigma^2, \xi)$ denote the skewed normal and skewed t distribution with ν degrees of freedom, respectively. ξ is the skewness parameter as described in Azzalini (2014). Higher values of ξ correspond to more skewed distributions.)

In these scenarios the true quantile functions are known. See Appendix B for an analytical derivation of the conditional quantiles of the multivariate t copula as well as the three-dimensional Clayton copula.

To assess the performance of each of the estimation methods we consider the estimated out-of-sample mean integrated square error ($\hat{\text{MISE}}_m$) of method m , defined as

$$\hat{\text{MISE}}_m := \frac{1}{R} \sum_{r=1}^R \left[\frac{1}{n_{eval}} \sum_{i=1}^{n_{eval}} \left\{ \hat{q}_{m,\alpha}^{(r)}(\mathbf{x}_{r,i}^{eval}) - q_\alpha(\mathbf{x}_{r,i}^{eval}) \right\}^2 \right], \quad (5.14)$$

where, for each replication $r = 1, \dots, R = 100$, we simulate a training data set $(y_{r,i}^{train}, \mathbf{x}_{r,i}^{train})$, $i = 1, \dots, n_{train}$, from the distribution of $(Y, \mathbf{X})'$, and further simulate an evaluation data set $(\mathbf{x}_{r,i}^{eval})$, $i = 1, \dots, n_{eval}$, of the distribution of \mathbf{X} . In this

study we choose $n_{eval} = \frac{1}{2}n_{train}$. For each replication r we estimate $\hat{q}_{m,\alpha}^{(r)}(\cdot)$, the quantile estimator of method m based on the training data set. Then, we use the evaluation data set to estimate the integrated square error between the resulting estimates and the true quantiles. Finally, the mean over all replications is taken to yield the estimated mean integrated square error.

In order to better compare the performance of the competitor methods relative to D-vine regression we also consider the relative estimated mean integrated square error of method m , defined as

$$\text{RMISE}_m := \frac{\hat{\text{MISE}}_m}{\hat{\text{MISE}}_{DVQR}}, \quad (5.15)$$

where $\hat{\text{MISE}}_{DVQR}$ is the estimated MISE of D-vine quantile regression. Values greater than one imply a worse relative performance compared to D-vine quantile regression. For each of the scenarios we present a table containing its results. The first two columns of each table specify the parameter settings for the margins and dependence structure. The third column gives information about the sample size of the training data set used to fit the models, while the fourth column determines which quantile is estimated. The fifth column contains the estimated MISE values obtained using D-vine quantile regression. The relative performances of the competitor methods can be found in columns six to nine in terms of the estimated RMISE as defined above.

Results for Scenario C3 At first we will consider the results of Scenario **C3** in [Table 3](#). A plot of the true quantile functions for this scenario is displayed in [Figure 5](#) of [Appendix B](#).

margins	δ	n_{train}	α	$\hat{\text{MISE}}_{DVQR}$	RMISE_m			
					LQR	BAQR	NPQR	SPQR
\mathcal{M}_1	δ_1	300	0.5	0.0118	4.08	2.87	4.39	0.97
			0.95	0.0252	3.94	3.13	4.23	1.09
		1000	0.5	0.0036	12.91	5.26	7.56	1.07
			0.95	0.0083	7.02	4.20	7.28	1.09
	δ_2	300	0.5	0.0029	10.60	5.72	9.24	1.63
			0.95	0.0171	7.21	4.21	6.80	1.22
		1000	0.5	0.0011	30.90	11.86	16.80	1.46
			0.95	0.0054	20.06	7.09	12.90	1.19
\mathcal{M}_2	δ_1	300	0.5	0.0078	6.13	3.71	6.74	1.12
			0.95	0.0776	3.40	2.43	6.74	1.06
		1000	0.5	0.0022	20.53	6.02	14.55	1.19
			0.95	0.0209	7.54	4.05	13.45	1.09
	δ_2	300	0.5	0.0036	10.55	6.59	19.24	1.98
			0.95	0.0664	3.29	2.43	8.76	1.17
		1000	0.5	0.0012	31.28	11.90	35.62	2.37
			0.95	0.0183	8.28	4.73	18.72	1.22

Table 3: $\hat{\text{MISE}}$ of D-vine quantile regression and the relative performances of the other estimation methods based on data generated by Scenario **C3** with different parameter settings as specified in [Table 2](#).

With regard to the estimated mean integrated square error of D-vine quantile regression (column 5) we see that the model fits very well with errors ranging in the order of magnitude 10^{-3} to 10^{-2} (we do not show plots of the fitted quantiles since they

are almost identical to those displayed in Figure 5). We generally observe that the 50%-quantile has a better fit than the 95%-quantile which is not surprising since the univariate estimator of the 50%-quantile is more robust. Further, it is clear that the prediction errors are reduced when the training sample size increases. It is also plausible that the size of errors drops when changing from the setting of low dependence (δ_1) to high dependence (δ_2). Concerning the marginal distributions it is reasonable that the performance of the 95%-quantile prediction worsens when the response distribution is skewed and heavy-tailed, since its marginal estimation is more imprecise, especially in the tails.

Comparing the estimated MISE values of D-vine quantile regression with the other methods described in Section 4 using the estimated relative mean integrated square error, we see that it outperforms its competitors by a great margin. The plots of the true conditional quantiles in Figure 5 imply that the linearity assumption of linear quantile regression is clearly violated in this scenario which results in severe prediction errors up to 30 times higher than our method. Using boosting additive models that allow for a nonlinear relationship manage to improve these results, however still lagging behind D-Vine quantile regression by a factor of 2.4 to 11.9. Further, in this example nonparametric quantile regression also seems too imprecise to be a serious competitor to D-vine quantile regression. Being conceptually closest to our method, it is not surprising that the semiparametric approach performs similarly well. However, except for one parameter setting it is still outperformed by the D-vine quantile regression considerably. Especially in the case of high dependence (δ_2) and skewed margins (\mathcal{M}_2) its predictions generate notably higher errors.

At this point we shortly focus on the computational times of the different quantile regression methods. Since they do not depend on the margins or copula parameter, we present in Table 4 the computational times for the exemplary setting δ_1 and \mathcal{M}_1 for different sizes of the training data set. The times are given in seconds needed for the $R=100$ repetitions of estimating the 0.5- and 0.95-quantiles.

n_{train}	DVQR	LQR	BAQR	NPQR	SPQR
300	446	0.79	44	3065	4125
1000	1364	1.23	84	22768	24144

Table 4: Computational times in seconds for the different quantile regression methods in the exemplary setting δ_1 and \mathcal{M}_1 .

Due to the simplicity of linear quantile regression its computation times are unbeatable. Also, boosting additive models are considerably faster than the remaining methods. However, we observe that D-vine quantile regression is much faster than its nonparametric and semiparametric competitors. Further, the computational time of D-vine quantile regression grows linearly in n_{train} while for the other two methods it seems to increase at a much higher rate. Moreover, we note that for the prediction of quantiles at different α -levels D-vine quantile regression has the advantage of having to fit the model only once and then being able to easily extract different conditional quantiles by evaluating the inverse h-functions. For the other methods a separate optimization has to be performed for each quantile level.

Results for Scenario t5 The competitive advantage of D-vine quantile regression over the competitor methods is also supported by the results of Scenario **t5** displayed

in Table 5. However, with the t-distribution being a little bit closer to the Gaussian

margins	δ	n_{train}	α	$\hat{\text{MISE}}_{DVQR}$	$\hat{\text{RMISE}}_m$			
					LQR	BAQR	NPQR	SPQR
\mathcal{M}_1	R_1	300	0.5	0.0357	0.62	1.28	2.68	1.46
			0.95	0.0774	1.98	1.65	2.26	1.17
		1000	0.5	0.0063	1.90	2.81	9.05	1.63
			0.95	0.0303	4.18	2.73	5.03	0.75
	R_2	300	0.5	0.0214	0.59	1.27	2.45	1.23
			0.95	0.0738	1.46	1.36	1.87	0.97
		1000	0.5	0.0052	1.34	2.33	6.56	2.18
			0.95	0.0278	3.44	2.58	3.80	0.98
\mathcal{M}_2	R_1	300	0.5	0.0319	0.91	1.92	3.81	1.60
			0.95	0.2912	1.37	1.08	1.64	1.04
		1000	0.5	0.0103	2.61	3.54	10.92	1.44
			0.95	0.1702	2.10	1.47	2.63	0.97
	R_2	300	0.5	0.0250	0.81	1.61	4.08	1.53
			0.95	0.1842	1.17	1.02	1.71	1.06
		1000	0.5	0.0078	1.90	2.64	12.74	2.13
			0.95	0.0948	2.05	1.45	3.27	1.17

Table 5: $\hat{\text{MISE}}$ of D-vine quantile regression and the relative performances of the other estimation methods based on data generated by Scenario **t5** with different parameter settings.

model, the linearity assumption of the linear quantile regression method is not violated as severely resulting in smaller relative errors. In fact, in the setting of a small training sample size $n_{train} = 300$, the prediction of the median is a bit more accurate than the prediction using any of the other methods. Nevertheless, in the tails of the distribution ($\alpha = 0.95$), where the t- and the Gaussian distribution differ the most, it loses this property. Again, the results of the boosting additive and nonparametric method are well behind the others with especially bad relative performance in predicting the median when the sample size is large. Concerning the prediction of median values D-vine quantile regression also outperforms the semiparametric approach whose relative errors lie between 1.23 and 2.18. In the tail they perform quite similar with relative errors ranging between 0.75 and 1.17. All in all, D-vine quantile regression is still the preferred method, especially for larger sample sizes.

Results for Scenario M5 Finally, Table 6 displays the results corresponding to Scenario **M5**. Motivated by Dette et al. (2014) who argue that a non-monotonic relationship between the response and the predictors cannot be modeled by a parametric copula, we consider Scenario **M5** as a case where no parametric D-vine (or R-vine) is able to perfectly capture the model's dependence structure. This becomes noticeable when looking at the estimated mean squared error of D-vine quantile regression which is clearly larger than in the previous examples. Also, we observe that the order of magnitude of the errors does not change when moving from a small to a large sample size implying a model bias. This model bias also seems to be visible when linear and semiparametric quantile regressions are used. All three models are quite close in performance with the linear model falling a bit behind. However, we observe that in this scenario nonparametric and boosting additive quantile regression methods excel.

σ	n_{train}	α	$\hat{\text{MISE}}_{DVQR}$	RMISE_m			
				LQR	BAQR	NPQR	SPQR
0.1	300	0.5	0.209	1.39	0.67	0.41	0.95
		0.95	0.514	1.37	0.68	0.30	1.04
	1000	0.5	0.228	1.42	0.61	0.34	0.95
		0.95	0.551	1.36	0.67	0.23	1.09
1	300	0.5	0.249	1.16	0.85	0.76	0.95
		0.95	0.312	1.19	0.90	1.02	0.92
	1000	0.5	0.215	1.40	0.75	0.61	0.95
		0.95	0.289	1.27	0.84	0.76	0.95

Table 6: $\hat{\text{MISE}}$ of D-vine quantile regression and the relative performances of the other estimation methods based on data generated by Scenario **M5** with different parameter settings.

Imposing no restrictions on the model assumptions, NPQR generates relative errors ranging between 0.23 and 1.02. This motivates further research to facilitate the inclusion of nonparametric pair-copulas (Nagler and Czado 2015) in the construction of the D-vine used for quantile regression to accommodate non-monotonicities. The need for this is also underlined in the setting $\sigma = 0.1$ and $n_{train} = 300$ by an exemplary scatterplot between Y and X_1 (which is the first variable chosen by the D-vine regression algorithm), displayed in the left panel of Figure 3 (together with a locally weighted scatterplot smoothing line).

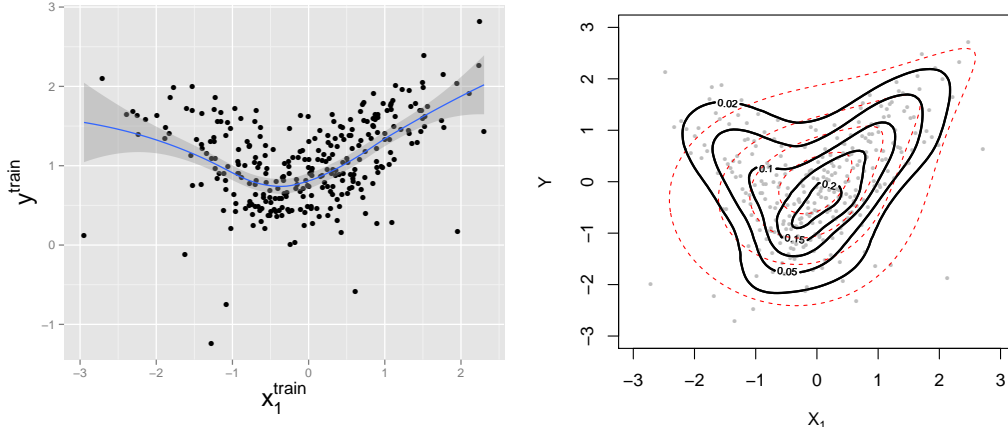


Figure 3: Left: Exemplary scatterplot between Y and X_1 in the setting $\sigma = 0.1$ and $n_{train} = 300$ of Scenario **M5**, together with a locally weighted scatterplot smoothing line. Right: Empirical nonparametric contour plot of the copula between X_1 and Y with standard normal margins (solid) and the corresponding contour plot of the fitted copula (dashed).

We clearly see a non-monotonic relationship between X_1 and Y . Currently, no pair-copula implemented in the `VineCopula` package is able to model such a dependence. The solid lines of the right panel of Figure 3 display the empirical nonparametric contour plot of the corresponding copula with standard normal margins generated by the function `kdecop` of the package `kdecopula` (Nagler 2015). In dashed lines the corresponding contour plot of the fitted parametric copula is shown (Joe copula with $\tau \approx 0.25$). The unsatisfying model fit is obvious and explains the rather high estimated MISE values. With this in mind, it is also understandable that a larger training sample

size does not help to improve the model fit and prediction accuracy of D-vine quantile regression for this example. However, we observe that the nonparametrically estimated copula manages to model the non-monotonic dependence of the data quite well (for visualization purposes we added the data points transformed to have standard normal margins as well). Hence, by using a nonparametric copula to model the dependence of the pair (Y, X_1) , a model misspecification as described in [Dette et al. \(2014\)](#) would be avoided. Further, as discussed in [Nagler and Czado \(2015\)](#) in detail, by modeling only bivariate copulas nonparametrically the dreaded curse of dimensionality is evaded.

6 Application to the CDS data set

As an application of D-vine quantile regression to real data we want to exploit interdependencies in the financial market in order to set up models that are able to forecast performances of single institutions as well as construct global stress tests. For these purposes we consider a data set containing 1371 daily observations (01/04/2006 – 10/25/2011) of log-returns of credit default swap (CDS) spreads with 5 year maturity of 38 European, US American and Asian-Pacific financial institutions in the banking and insurance sectors. This data set has already been analyzed by [Brechmann et al. \(2013\)](#), who argue that CDS spreads are a viable and accurate measure of a company’s creditworthiness. After applying an appropriate GARCH model to each of the univariate time series in order to get approximately i.i.d. residuals, they perform stress tests. By sampling from a conditional C-vine they stress one company at a time (i.e. setting it to its 90%/95%/99% quantile) and examine the impact on the other institutions conditioned on this stress event. With our method we can go even further and consider scenarios where more than one company is in distress. This allows us to investigate the spillover effects of a financial crisis in a certain region or branch to other regions and branches. Additional to the stress tests we use D-vine quantile regression models later in this chapter in order to predict an institution’s CDS spread log-returns given its past values and the log-returns of peer companies. In an out-of-sample test we underline the superiority of our approach once again.

6.1 Stress testing

The financial institutions considered in the stress test are 18 banks and 20 (re-)insurers from the regions USA, Europe and Asia-Pacific:

- **Banks:** *USA* (Goldman Sachs (GS), JP Morgan Chase (JPM), Citigroup), *Europe* (Banco Bilbao Vizcaya Argentaria (BBVA), Banco Santander (BS), Barclays, BNP Paribas, Deutsche Bank (DB), Intesa Sanpaolo, Royal Bank of Scotland (RBS), Société Générale (SG), Standard Chartered (StanCha), UBS, Unicredit), *Asia-Pacific* (Bank of China (BoC), Kookmin Bank, Sumitomo Mitsui, Westpac Banking)
- **(Re-)Insurers:** *USA* (ACE, Allstate, American International Group (AIG), Chubb, Hartford Financial Services, XL Group), *Europe* (Aegon, Allianz, Assicurazioni Generali, Aviva, AXA, Hannover Rück (HR), Legal & General (LG), Munich Re (MR), Prudential, SCOR, Swiss Re (SR), Zurich Insurance), *Asia-Pacific* (Tokio Marine (TM), QBE Insurance)

We consider three stress scenarios, corresponding to crises originating in different sectors, and investigate the resulting spillover effects. For this, we proceed as in [Brechmann et al. \(2013\)](#) and remove serial dependencies from each of the 38 univariate time series by fitting adequate GARCH models. The resulting residuals r_j , $j = 1, \dots, 38$, which are approximately independent and distributed according to their model's estimated innovations distribution \hat{F}_j , carry the information about the dependence structure between the institutions. We consider company j to be stressed at level $\kappa \in (0, 1)$, if its residual r_j takes on the $100\kappa\%$ -quantile of its innovation distribution \hat{F}_j , i.e. $r_j = \hat{F}_j^{-1}(\kappa)$. This is equivalent to the PIT transformed variable $u_j := \hat{F}_j(r_j)$ taking on value κ . Likewise, we are interested in the resulting predicted quantile levels of the non-stressed companies. This allows us to directly work on the u-scale and consider the PIT transformed variables $u_j := \hat{F}_j(r_j)$, $j = 1, \dots, 38$, and their dependencies.

In Scenario 1, we analyze the effect of stressing the European *systemic* banks as specified by [International Monetary Fund \(2009\)](#) at different stress levels. Therefore, we stress the banks Banco Santander, Barclays, BNP Paribas, Deutsche Bank, Royal Bank of Scotland, Société Générale, UBS and Unicredit at level $\kappa \in \{0.9, 0.95, 0.99\}$ (corresponding to moderate, severe and extreme stress scenarios) and use D-vine quantile regression to estimate the conditional medians of the remaining institutions conditioned on this stress event. This way, we can assess the spillover effect to other sectors and regions. The left panel of [Figure 4](#) shows the results of the stress test of Scenario 1. For each institution the predicted median values for the three stress levels are coded by circles (moderate stress with $\kappa = 0.9$), diamonds (severe stress with $\kappa = 0.95$) and triangles (extreme stress with $\kappa = 0.99$). For visualization, the currently stressed institutions' names are printed in ***bold and italic***. Further, solid lines separate the geographical regions (Europe in the upper, USA in the middle and Asia-Pacific in the lower panel), while dashed lines separate banks (upper) from insurance companies (lower).

We observe that the spillover effect is strongest for European insurances with predicted median values of up to 0.98 (Allianz and Aviva) for the extreme stress scenario. The comparably small values of the British bank Standard Chartered are explained by the fact that according to their annual report of 2014¹ 90% of the bank's income and profits are earned in Asia, Africa and the Middle East. Similar arguments holds for the British insurance company L&G with operations in Asia and the United States². Another group that is affected quite strongly by this stress scenario are the US banks with predicted median values exceeding 0.9 in the extreme stress case. However, we observe that the geographic spillover effect is stronger than the institutional one, because the effect on European insurance companies is stronger than the effect on US banks. US insurance companies as well as the Asian-Pacific market are also affected by the stress scenario, but not as severe as the other groups.

It is interesting to see that in Scenario 2, where we only stress the three major European banks Barclays, BNP Paribas and Deutsche Bank, the results of the stress test are very similar to those of Scenario 1. We can conclude, that for a crisis to evolve it suffices that only few but important banks default. In the last scenario, we analyze the spillover effect of a default of the US American banking system. Therefore, we stress the banks Citigroup, Goldman Sachs and JP Morgan Chase at level $\kappa \in \{0.9, 0.95, 0.99\}$ and estimate the conditional medians of the remaining institutions conditioned on this stress event. Again, we see the quite strong interconnectedness between US banks

¹https://www.sc.com/annual-report/2014/documents/SCB_ARA_2014_full_report.pdf

²<http://www.legalandgeneralgroup.com/all-our-sites/>

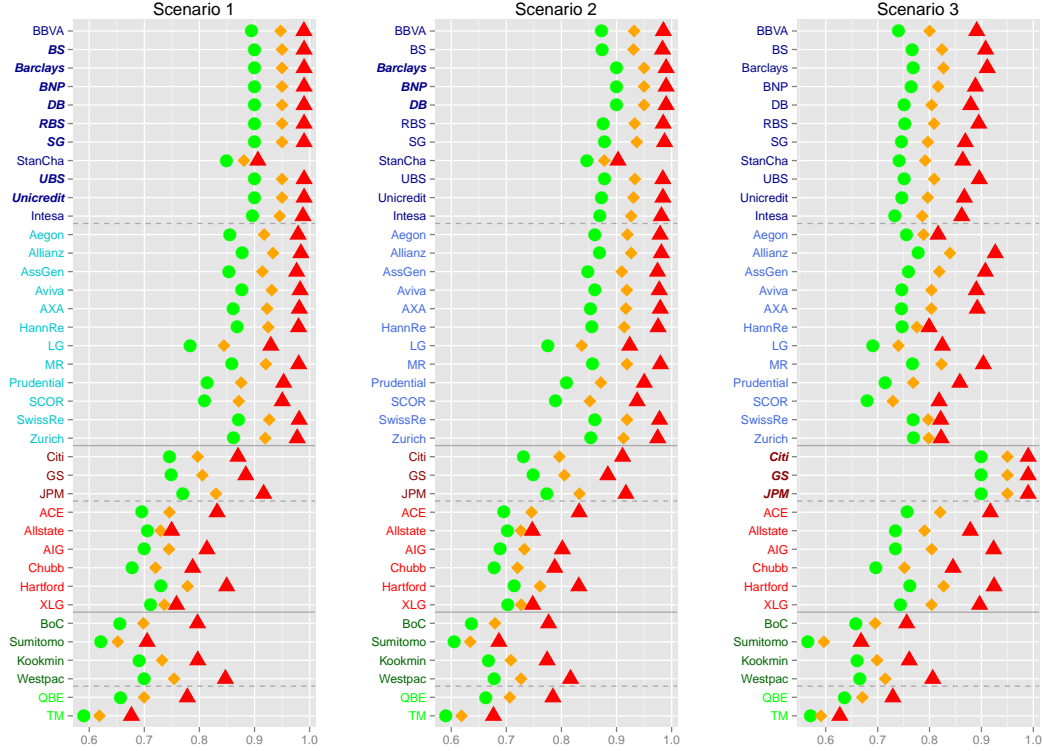


Figure 4: Stress Tests stressing European systemic banks (left), major European banks (middle) and US banks (right). For each institution the predicted median values for the three stress levels are represented by circles (moderate stress with $\kappa = 0.9$), diamonds (severe stress with $\kappa = 0.95$) and triangles (extreme stress with $\kappa = 0.99$).

and insurance companies, as well as between US banks and the European market, and observe rather weak spillover effects on the Asian-Pacific sector.

6.2 Out-of-sample quantile prediction

In order to assess the prediction performance of D-vine quantile regression we conduct an out-of-sample test. For this, we consider the raw CDS log-return data (i.e. no time series models were fitted). The goal is to predict the quantiles of Y_t , i.e. the CDS log-returns of company Y at time t , given its past values Y_{t-1}, \dots, Y_{t-l} (for some lag l) and the past and current values of its peer companies $\mathbf{X}_t, \dots, \mathbf{X}_{t-l}$. In the out-of-sample test, we split the data set at a certain time point into a training data set, based on which we will fit the different quantile regression models, and an evaluation data set used to assess the performance of the methods. As a split point we choose the date November 8th, 2010, such that the training data set contains $n_{train} = 1121$ observations and the evaluation data set contains $n_{eval} = 250$ observations. Note that this sample is not i.i.d. However, since we can assume that it fulfills the strong mixing condition due to low autocorrelations for lags higher than 3 (Hansen 2008), the approach of estimating the marginals nonparametrically is justified.

Predicting CDS log-returns of Goldman Sachs We consider the CDS log-returns of the US American bank Goldman Sachs as response. It is not possible to use all of the remaining 37 institutions as predictors since the computational times for the nonpara-

metric and semiparametric quantile regression methods would be too high. Therefore, we reduce the number of predictors by running the D-vine regression algorithm on all institutions with lag value l equal to 2 and choosing only the meaningful institutions as covariates. These are US American banks JP Morgan Chase and Citigroup as well as the US insurance company Allstate and the British Bank Barclays. The highest lag value chosen was 1.

Applying the D-vine quantile regression algorithm to the data set we obtain a D-vine with order $GS - JPM - Citi - Allstate - Barclays - GS_{t-1}$ (a list of chosen pair-copula families with corresponding parameters can be found in [Appendix C](#)). Not surprisingly, the algorithm ranks fellow US American Banks JP Morgan Chase and Citigroup as most influential. Interestingly, the autocorrelation term GS_{t-1} appears in last place, while the lagged values of the peer companies are even regarded as unimportant. For each of the four quantile regression methods we present five different measures assessing the performance of quantile prediction in [Table 7](#). In the first row, we compare the predicted medians $\hat{q}_{0.5,i} := \hat{q}_{0.5}(Y_{i-1}, \mathbf{X}_i, \mathbf{X}_{i-1})$ with the observations Y_i^{eval} , $i = 1, \dots, n_{eval}$ using the estimated mean squared error: $\hat{MSE} := n_{eval}^{-1} \sum_{i=1}^{n_{eval}} (Y_i^{eval} - \hat{q}_{0.5,i})^2$. Further, the second and third rows of the table concern the 95% prediction interval of observation i , denoted by $PI_{0.95,i} := (\hat{q}_{0.025,i}, \hat{q}_{0.975,i})$. The second row gives the coverage percentage (CP) of how often the observation lies in the prediction interval: $CP_{PI_{0.95}} := n_{eval}^{-1} \sum_{i=1}^{n_{eval}} \mathbb{1}\{Y_i^{eval} \in PI_{0.95,i}\}$, where $\mathbb{1}\{\cdot\}$ denotes the indicator function. We expect a good model to have a coverage percentage close to 95% since that is the coverage percentage of the true model. Another interesting measure to consider is the mean length of the prediction interval, presented in the third row of the table, defined as $ML(PI_{0.95}) := n_{eval}^{-1} \sum_{i=1}^{n_{eval}} (\hat{q}_{0.975,i} - \hat{q}_{0.025,i})$. Finally, in the last two rows of the table we display the coverage percentages according to the predicted 95% and 99.5% quantiles, i.e. the percentages of how often the observation is smaller than $\hat{q}_{0.95,i}$ and $\hat{q}_{0.995,i}$, respectively. This is motivated by the financial risk measure *Value-at-Risk*, which is defined as a certain quantile of a loss distribution. Hence, we define $CP_{VaR_\alpha} := n_{eval}^{-1} \sum_{i=1}^{n_{eval}} \mathbb{1}\{Y_i^{eval} < \hat{q}_{\alpha,i}\}$ as a measure of how well the VaR_α is predicted.

Method	DVQR	LQR	BAQR	NPQR	SPQR
$\hat{MSE} \times 1000$	0.680	0.689	0.715	1.008	0.728
$CP_{PI_{0.95}}$	0.948	0.972	0.956	0.968	0.956
$ML(PI_{0.95})$	0.098	0.145	0.110	0.122	0.093
$CP_{VaR_{0.95}}$	0.940	0.960	0.940	0.956	0.932
$CP_{VaR_{0.995}}$	0.996	0.988	0.984	1.000	0.988

Table 7: Results of the out-of-sample test predicting the CDS log-returns of Goldman Sachs using its past values and the current returns of JP Morgan Chase, Citigroup, Allstate and Barclays using different conditional quantile estimation methods.

Examining [Table 7](#) we observe that D-vine quantile regression has the lowest mean squared error. Also, the coverage percentages of the prediction interval and the two VaR values are very close to the true values of 0.95, 0.95 and 0.995, respectively, while the average length of the prediction interval is comparatively small. Looking at the competitor quantile regression methods we see that, while linear quantile regression manages to perform quite well at predicting the conditional median, it fails to estimate the more extreme quantiles lying in the tails of the distribution. On the one hand, the estimated 95% prediction interval is on average almost 1.5 times as large as the one of our method, causing the coverage percentage to be too high. On the other hand, the

prediction of the 99.5% quantile is too low resulting in more observed exceedances than predicted by this linear model. Interestingly, the performance of the nonparametric estimator is by far the worst. Only its prediction of the 95% quantile seems reasonable. Further, for the given data set copula-based quantile regression is worse than our method when it comes to predicting the median and the 95% and 99.5% quantiles. However, it manages to attain the smallest mean length of the 95% confidence interval while still having an adequate coverage percentage of around 95%. Lastly, boosting additive quantile regression achieves acceptable results that are still outperformed by D-vine quantile regression. All in all, this real data application has supported the findings of the simulation study of [Section 5](#), that D-vine quantile regression is the method to choose if you want to estimate conditional quantiles.

7 Conclusions and further research

A new method to predict conditional quantiles is proposed. We have seen that the usage of the flexible D-vine class facilitates fast and accurate estimation. Analyzing the log returns of the CDS spreads of international banks and insurances we extend the analysis of [Brechmann et al. \(2013\)](#). While they analyze the spillover effects stressing only one institution by simulating from a conditioned C-vine, with our new method we are able to perform stress tests that are conditioned on multiple banks and insurances being in distress. In our analysis we found out that the spillover effect is mainly driven by geography, so that European banks have a greater influence on European insurances than on US American banks. Further, the claim of [Brechmann et al. \(2013\)](#) that US banks have a stronger influence on the international financial market than European banks is not supported by our analysis. Stressing the major European banks had a greater overall impact on the financial system than stressing the US American banks. Moreover, we have seen that there is still room for future research. With the bivariate pair-copulas implemented in the package `VineCopula` there is currently no possibility to model non-monotonic dependencies between a response and its predictors as is already pointed out in [Dette et al. \(2014\)](#). A remedy for this is the inclusion of nonparametric pair-copulas in the construction of the D-vine used for quantile regression. A further topic of ongoing research is the implementation of D-vines with mixed discrete and continuous margins as discussed in [Stöber et al. \(2013\)](#) in order to yield quantile regression models allowing for discrete as well as continuous responses and predictors.

Appendix

A D-vine regression algorithm

Algorithm A.1 (D-vine regression algorithm).

Input: pseudo copula data $\mathbf{v} = (v^{(i)})_{i=1,\dots,n}$, $\mathcal{U} = (u_j^{(i)})_{j=1,\dots,d, i=1,\dots,n}$

Initialize global variables:

- 1: $I_0 \leftarrow \{1, \dots, d\}$
- 2: `global.max.cll` $\leftarrow -\infty$

Step 1:


```

3: for  $j \in I_0$  do
4:   Use BiCopSelect (Algorithm A.2) to estimate  $C_{V,U_j}$ , the AIC-optimal pair-
     copula for  $(V, U_j)'$  based on the data  $(\mathbf{v}, \mathbf{u}_j)$ 
5:    $\text{ccl}_j \leftarrow \text{ccl}(\hat{c}_{V,U_j}; \mathbf{v}, \mathbf{u}_j)$ 
6: end for
7:  $l_1 \leftarrow \arg\max_{j \in I_0} \text{ccl}_j$ 
   Update global variables:
8:  $\text{global.max.ccl} \leftarrow \max_{j \in I_0} \text{ccl}_j$ 
9:  $I_1 \leftarrow I_0 \setminus l_1$ 

```

Step k : (the first tree of the current D-vine with $k - 1$ covariates has the order $V-U_{l_1}-\dots-U_{l_{k-1}}$ (see the black graph in Figure 2). I_{k-1} contains the indices of the covariates that have not yet been chosen for the model and global.max.ccl is equal to the conditional log-likelihood of the current model)

```

10: for  $j \in I_{k-1}$  do
11:   Estimate  $C_{V,U_j|U_{l_1},\dots,U_{l_{k-1}}}$  and  $C_{U_{l_i},U_j|U_{l_{i+1}},\dots,U_{l_{k-1}}}$ ,  $i = 1, \dots, k-1$ , i.e. the pair-
     copulas needed to extend the current D-vine to a D-vine with order  $V-U_{l_1}-\dots-U_{l_{k-1}}-U_j$  (compare Figure 2)
12:    $\text{ccl}_j \leftarrow \text{ccl}(\hat{c}_{V,U_{l_1},\dots,U_{l_{k-1}},U_j}; \mathbf{v}, \mathbf{u}_{l_1}, \dots, \mathbf{u}_{l_{k-1}}, \mathbf{u}_j)$ 
13: end for
14: if  $\max_{j \in I_{k-1}} \text{ccl}_j < \text{global.max.ccl}$  then
15:   return D-vine with order  $V-U_{l_1}-\dots-U_{l_{k-1}}$ 
16: end if
17:  $\text{global.max.ccl} \leftarrow \max_{j \in I_0} \text{ccl}_j$ 
18:  $l_k \leftarrow \arg\max_{j \in I_{k-1}} \text{ccl}_j$ 
19:  $I_k \leftarrow I_{k-1} \setminus l_k$ 

```

20: After step d has been exercised, **return** D-vine with order $V-U_{l_1}-\dots-U_{l_d}$

In lines 5 and 12 of the algorithm, the AIC- and BIC-corrected conditional log-likelihoods may be used instead of the regular ccl. In this case all maximum functions have to be replaced by minimum functions.

The algorithm repeatedly uses the function BiCopSelect of the package VineCopula, which fits the BIC-optimal pair-copula to copula data after testing for independence (for details see Genest and Favre (2007)):

Algorithm A.2 (BiCopSelect).

Input: i.i.d. copula data $(u_1^{(i)}, u_2^{(i)})$, $i = 1, \dots, n$; significance level for independence test β .

```

1: Calculate the empirical Kendall's Tau  $\hat{\tau}$  and the test statistic  $t = \sqrt{\frac{9n(n-1)}{2(2n+5)}} |\hat{\tau}|$ .
2: if  $2(1 - \Phi(t)) \geq \beta$  then
3:   Return the independence copula  $c_\pi$ .
4: end if
5: for each pair-copula family do
6:    $\theta_{ML} = \arg\max_{\theta \in \Theta} \prod_{i=1}^n c(u_1^{(i)}, u_2^{(i)}; \theta)$ .
7:    $\text{BIC} = -2 \sum_{i=1}^n \ln(c(u_1^{(i)}, u_2^{(i)}; \theta_{ML})) + \ln(n) |\theta_{ML}|$ .
8: end for
9: Return the copula with the lowest BIC-value with its MLE parameter  $\theta_{ML}$ .

```

Here, $|\theta_{ML}|$ denotes the number of parameters of the copula family under consideration.

Regarding the copula families in line 5 of the code, the pair-copula families currently implemented in the `VineCopula` package are Gaussian, t^* , Clayton, Gumbel, Frank, Joe, BB1*, BB6*, BB7*, BB8* and Tawn* with respective rotations. Families marked with a star are two-parametric, while the others are one-parametric.

B Analytically derived conditional copula quantiles

B.1 Conditional copula quantile function based on the Gaussian copula

It is a well known fact that the conditional distribution of a multivariate normal distributed random vector again is normally distributed with shifted mean and covariance matrix ([Brachinger et al. \(1996\)](#)). Therefore, assuming

$$(Y, X_1, \dots, X_d) \sim N \left((\mu_Y, \mu_{\mathbf{X}})', \begin{pmatrix} \sigma_Y^2 & \Sigma_{Y,\mathbf{X}} \\ \Sigma'_{Y,\mathbf{X}} & \Sigma_{\mathbf{X},\mathbf{X}} \end{pmatrix} \right), \quad (\text{B.16})$$

the conditional quantile of Y at level $\alpha \in (0, 1)$ given $\mathbf{X} = \mathbf{x}$ can be calculated as

$$F_{Y|\mathbf{X}}^{-1}(\alpha|\mathbf{x}) = \Phi^{-1}(\alpha; \mu_c(\mathbf{x}), \sigma_c^2),$$

where $\Phi(\cdot; \mu, \sigma^2)$ denotes the distribution function of a $N(\mu, \sigma^2)$ -distributed random variable and the conditional parameters $\mu_c(\mathbf{x})$ and σ_c^2 are given as

$$\mu_c(\mathbf{x}) := \mu_Y + \Sigma_{Y,\mathbf{X}} \Sigma_{\mathbf{X},\mathbf{X}}^{-1} (\mathbf{x} - \mu_{\mathbf{X}}),$$

and

$$\Sigma_c := \sigma_Y^2 - \Sigma_{Y,\mathbf{X}} \Sigma_{\mathbf{X},\mathbf{X}}^{-1} \Sigma'_{Y,\mathbf{X}}.$$

Hence, by inverting [Equation \(3.5\)](#) the conditional quantile of the Gaussian copula is given as

$$C_{V|U_1, \dots, U_d}^{-1}(\alpha|u_1, \dots, u_d) = \Phi(\Phi^{-1}(\alpha; \mu_c(\mathbf{x}), \sigma_c^2); \mu_Y, \sigma_Y^2),$$

where $\mathbf{x} = (\Phi^{-1}(u_1; \mu_1, \sigma_1^2), \dots, \Phi^{-1}(u_d; \mu_d, \sigma_d^2))'$.

B.2 Conditional copula quantile function based on the Student's t copula

Similarly, following [Kotz and Nadarajah \(2004\)](#), we know that the conditional distribution of a multivariate t distributed random vector with ν degrees of freedom again is t distributed with shifted mean, covariance matrix and degrees of freedom. Therefore, with the same partitions of the mean and covariance matrix as in [Equation \(B.16\)](#), the conditional quantile of Y at level $\alpha \in (0, 1)$ given $\mathbf{X} = \mathbf{x}$ can be calculated as

$$F_{Y|\mathbf{X}}^{-1}(\alpha|\mathbf{x}) = t^{-1}(\alpha; \nu_c, \mu_c(\mathbf{x}), \sigma_c^2(\mathbf{x})),$$

where $t(\cdot; \nu, \mu, \sigma^2)$ denotes the distribution function of a $t(\nu, \mu, \sigma^2)$ -distributed random variable and the conditional parameters ν_c , $\mu_c(\mathbf{x})$ and $\sigma_c^2(\mathbf{x})$ are given as

$$\nu_c := \nu + d,$$

$$\mu_c(\mathbf{x}) := \mu_Y + \Sigma_{Y,\mathbf{X}} \Sigma_{\mathbf{X},\mathbf{X}}^{-1} (\mathbf{x} - \mu_{\mathbf{X}}),$$

and

$$\sigma_c(\mathbf{x}) := \frac{\nu + (\mathbf{x} - \mu_{\mathbf{X}})' \Sigma_{\mathbf{X},\mathbf{X}}^{-1} (\mathbf{x} - \mu_{\mathbf{X}})}{\nu + d} (\sigma_Y^2 - \Sigma_{Y,\mathbf{X}} \Sigma_{\mathbf{X},\mathbf{X}}^{-1} \Sigma_{Y,\mathbf{X}}').$$

Consequently, the conditional quantile of the t copula is given as

$$C_{V|U_1, \dots, U_d}^{-1}(\alpha | u_1, \dots, u_d) = t(t^{-1}(\alpha; \nu_c, \mu_c(\mathbf{x}), \sigma_c^2(\mathbf{x})); \nu, \mu_Y, \sigma_Y^2),$$

where $\mathbf{x} = (t^{-1}(u_1; \nu, \mu_1, \sigma_1^2), \dots, t^{-1}(u_d; \nu, \mu_d, \sigma_d^2))'$.

B.3 Conditional copula quantile function based on the two- and three-dimensional Clayton copula

For the bivariate Clayton copula with parameter $\delta > 0$ ($C_{U,V}(u, v) = (u^{-\delta} + v^{-\delta} - 1)^{-\frac{1}{\delta}}$) the conditional quantile can easily be derived by inverting the h-function, yielding

$$C_{V|U}^{-1}(\alpha | u) = \left\{ \left(\alpha^{-\frac{\delta}{1+\delta}} - 1 \right) u^{-\delta} + 1 \right\}^{-\frac{1}{\delta}}.$$

The conditional quantile function of a three-dimensional Clayton copula with parameter $\delta > 0$ ($C_{U,V,W}(u, v, w) = (u^{-\delta} + v^{-\delta} + w^{-\delta} - 2)^{-\frac{1}{\delta}}$) is given by

$$C_{U|V,W}^{-1}(\alpha | v, w) = \left\{ (\alpha^{-\frac{\delta}{1+2\delta}} - 1)(v^{-\delta} + w^{-\delta} - 1) + 1 \right\}^{-\frac{1}{\delta}}.$$

To derive this, let (U, V, W) follow a three-dimensional Clayton copula with parameter $\delta > 0$. Then,

$$C_{U,V,W}(u, v, w) = (u^{-\delta} + v^{-\delta} + w^{-\delta} - 2)^{-\frac{1}{\delta}}.$$

In order to derive the conditional distribution of U given $V = v$ and $W = w$ we use a version of [Equation \(2.3\)](#):

$$C_{U|V,W}(u|v, w) = \frac{\partial}{\partial \xi} C_{U,V,W}(C_{U|W}(u|w), \xi; w) \Big|_{\xi=C_{V|W}(v|w)} \quad (\text{B.17})$$

For the derivation of $C_{U|W}(u|w)$ we use the fact that $C_{U,W}(u, w) = C_{U,V,W}(u, 1, w) = (u^{-\delta} + w^{-\delta} - 1)^{-\frac{1}{\delta}}$. Hence,

$$\begin{aligned} C_{U|W}(u|w) &= \frac{\partial}{\partial w} C_{U,W}(u, w) = \frac{\partial}{\partial w} (u^{-\delta} + w^{-\delta} - 1)^{-\frac{1}{\delta}} \\ &= (u^{-\delta} + w^{-\delta} - 1)^{-\frac{1}{\delta}-1} w^{-\delta-1}, \end{aligned}$$

and similarly,

$$C_{V|W}(v|w) = (v^{-\delta} + w^{-\delta} - 1)^{-\frac{1}{\delta}-1} w^{-\delta-1}.$$

The corresponding inverse functions are given by

$$C_{U|W}^{-1}(u|w) = \left\{ w^{-\delta} \left(u^{-\frac{\delta}{1+\delta}} - 1 \right) + 1 \right\}^{-\frac{1}{\delta}}$$

and

$$C_{V|W}^{-1}(v|w) = \left\{ w^{-\delta} \left(v^{-\frac{\delta}{1+\delta}} - 1 \right) + 1 \right\}^{-\frac{1}{\delta}}.$$

Further,

$$C_{U,V|W}(u, v|w) = \frac{\partial}{\partial w} C_{U,V,W}(u, v, w) = (u^{-\delta} + v^{-\delta} + w^{-\delta} - 2)^{-\frac{1}{\delta}-1} w^{-\delta-1},$$

and therefore,

$$\begin{aligned} C_{U,V;W}(u, v; w) &= C_{U,V|W} \left(C_{U|W}^{-1}(u|w), C_{V|W}^{-1}(v|w)|w \right) \\ &= \left\{ w^{-\delta} \left(u^{-\frac{\delta}{1+\delta}} - 1 \right) + 1 + w^{-\delta} \left(v^{-\frac{\delta}{1+\delta}} - 1 \right) + 1 + w^{-\delta} - 2 \right\}^{-\frac{1}{\delta}-1} w^{-\delta-1} \\ &= \left\{ w^{-\delta} \left(u^{-\frac{\delta}{1+\delta}} + v^{-\frac{\delta}{1+\delta}} - 1 \right) \right\}^{-\frac{1+\delta}{\delta}} w^{-\delta-1} = \left(u^{-\frac{\delta}{1+\delta}} + v^{-\frac{\delta}{1+\delta}} - 1 \right)^{-\frac{1+\delta}{\delta}}, \end{aligned}$$

which is independent of w and is the distribution function of a bivariate Clayton copula with parameter $\frac{\delta}{1+\delta}$.

Now it remains to calculate the differential in [Equation \(B.17\)](#) and plug in the conditional distributions:

$$\frac{\partial}{\partial \xi_2} C_{U,V;W}(\xi_1, \xi_2; w) = \left(\xi_1^{-\frac{\delta}{1+\delta}} + \xi_2^{-\frac{\delta}{1+\delta}} - 1 \right)^{-\frac{1+\delta}{\delta}-1} \xi_2^{-\frac{\delta}{1+\delta}-1}.$$

Finally,

$$\begin{aligned} C_{U|V,W}(u|v, w) &= \left\{ (v^{-\delta} + w^{-\delta} - 1) w^{\delta} + (u^{-\delta} + w^{-\delta} - 1) w^{\delta} - 1 \right\}^{-\frac{1+2\delta}{\delta}} \\ &\quad \cdot (v^{-\delta} + w^{-\delta} - 1)^{\frac{1+\delta}{\delta} \cdot \frac{1+2\delta}{1+\delta}} w^{(-\delta-1) \cdot \frac{-1-2\delta}{1+\delta}} \\ &= w^{-(1+2\delta)} (v^{-\delta} + w^{-\delta} - 1 + u^{-\delta} + w^{-\delta} - 1 - w^{-\delta})^{-\frac{1+2\delta}{\delta}} \\ &\quad \cdot (v^{-\delta} + w^{-\delta} - 1)^{\frac{1+2\delta}{\delta}} w^{1+2\delta} \\ &= (u^{-\delta} + v^{-\delta} + w^{-\delta} - 2)^{-\frac{1+2\delta}{\delta}} \cdot (v^{-\delta} + w^{-\delta} - 1)^{\frac{1+2\delta}{\delta}} \end{aligned}$$

Inversion with respect to u gives us the sought after conditional quantile function of a three-dimensional Clayton copula:

$$C_{U|V,W}^{-1}(\alpha|v, w) = \left\{ (\alpha^{-\frac{\delta}{1+2\delta}} - 1)(v^{-\delta} + w^{-\delta} - 1) + 1 \right\}^{-\frac{1}{\delta}}.$$

A visualization of the conditional median and 95% quantile function of a three-dimensional distribution with underlying Clayton copula with parameter and marginals as specified in [Table 2](#) is given in [Figure 5](#).

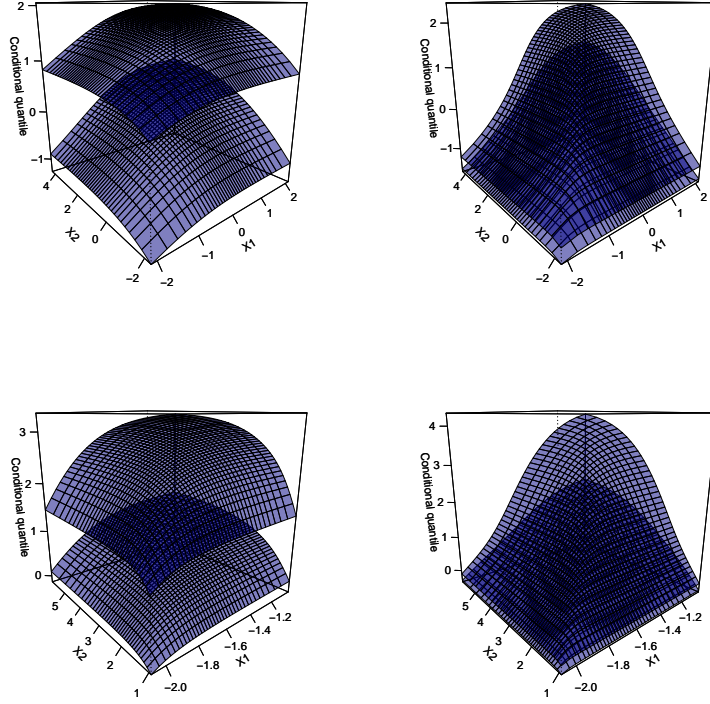


Figure 5: Conditional median and 95% quantile function of a three-dimensional distribution with underlying Clayton copula with parameter δ_1 (left panels), δ_2 (right panels), and marginals \mathcal{M}_1 (upper panels), \mathcal{M}_2 (lower panels) as specified in Table 2.

C Specifications of the D-vines used for the out-of-sample quantile prediction

Copula	family	τ	ν
C_{12}	t	0.49	2.9
C_{23}	t	0.46	2.7
C_{34}	t	0.23	4.3
C_{45}	t	0.22	4.1
C_{56}	t	0.14	3.6
$C_{13;2}$	t	0.23	4.7
$C_{24;3}$	SC	0.09	
$C_{35;4}$	t	0.24	4.6
$C_{46;5}$	t	0.12	10.9
$C_{14;23}$	t	0.13	5.9
$C_{25;34}$	t	0.16	6.2
$C_{36;45}$	SC	0.10	
$C_{15;234}$	t	0.12	10.0
$C_{26;345}$	SG	0.04	
$C_{16;2345}$	t	0.04	7.2

Here, $1 \leftrightarrow \text{GS}$, $2 \leftrightarrow \text{JPM}$, $3 \leftrightarrow \text{Citi}$, $4 \leftrightarrow \text{Allstate}$, $5 \leftrightarrow \text{Barclays}$, $6 \leftrightarrow \text{GS}_t - 1$. Regarding pair-copula families, t denotes the bivariate t copula with ν degrees of freedom, SC the bivariate survival Clayton copula and SG the bivariate survival Gumbel copula.

Acknowledgment

The second author is supported by the German Research Foundation (DFG grant GZ 86/4-1).

References

- Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009), “Pair-copula constructions of multiple dependence,” *Insurance: Mathematics and economics*, 44, 182–198.
- Adrian, T. and Brunnermeier, M. K. (2011), “CoVaR,” Tech. rep., National Bureau of Economic Research.
- Azzalini, A with the collaboration of Capitanio, A. (2014), *The Skew-Normal and Related Families*, Cambridge University Press, IMS Monographs series.
- Bedford, T. and Cooke, R. M. (2002), “Vines: A new graphical model for dependent random variables,” *Annals of Statistics*, 1031–1068.
- Bernard, C. and Czado, C. (2015), “Conditional quantiles and tail dependence,” *Journal of Multivariate Analysis*, 138, 104–126.
- Bouyé, E. and Salmon, M. (2009), “Dynamic copula quantile regressions and tail area dynamic dependence in Forex markets,” *The European Journal of Finance*, 15, 721–750.
- Brachinger, W., Fahrmeir, L., Hamerle, A., and Tutz, G. (1996), *Multivariate statistische verfahren*, Walter de Gruyter.
- Brechmann, E. C., Hendrich, K., and Czado, C. (2013), “Conditional copula simulation for systemic risk stress testing,” *Insurance: Mathematics and Economics*, 53, 722–732.
- Brownlees, C. T. and Engle, R. F. (2012), “Volatility, correlation and tails for systemic risk measurement,” *Available at SSRN 1611229*.
- Chen, X., Koenker, R., and Xiao, Z. (2009), “Copula-based nonlinear quantile autoregression,” *The Econometrics Journal*, 12, S50–S67.
- Cooke, R. M., Joe, H., and Chang, B. (2015), “Vine Regression,” *Resources for the Future Discussion Paper*, 15–52.
- Czado, C. (2010), “Pair-copula constructions of multivariate copulas,” in *Copula theory and its applications*, Springer, pp. 93–109.
- Detle, H., Van Hecke, R., and Volgushev, S. (2014), “Some comments on copula-based regression,” *Journal of the American Statistical Association*, 109, 1319–1324.
- Duong, T. (2015a), *ks: Kernel Smoothing*, r package version 1.9.4.
- (2015b), “Non-parametric smoothed estimation of multivariate cumulative distribution and survival functions, and receiver operating characteristic curves,” *Journal of the Korean Statistical Society*.

- Fenske, N., Kneib, T., and Hothorn, T. (2012), “Identifying risk factors for severe childhood malnutrition by boosting additive quantile regression,” *Journal of the American Statistical Association*.
- Genest, C. and Favre, A.-C. (2007), “Everything you always wanted to know about copula modeling but were afraid to ask,” *Journal of hydrologic engineering*, 12, 347–368.
- Haff, I. H., Aas, K., and Frigessi, A. (2010), “On the simplified pair-copula construction – Simply useful or too simplistic?” *Journal of Multivariate Analysis*, 101, 1296–1310.
- Hansen, B. E. (2008), “Uniform convergence rates for kernel estimation with dependent data,” *Econometric Theory*, 24, 726–748.
- Hayfield, T. and Racine, J. S. (2008), “Nonparametric Econometrics: The np Package,” *Journal of Statistical Software*, 27.
- Hothorn, T., Buehlmann, P., Kneib, T., Schmid, M., and Hofner, B. (2015), *mboost: Model-Based Boosting*, R package version R package version 2.5-0.
- Hwang, C. and Shim, J. (2005), “A simple quantile regression via support vector machine,” in *Advances in Natural Computation*, Springer, pp. 512–520.
- International Monetary Fund (2009), *Global Financial Stability Report*, Citeseer.
- Joe, H. (1997), *Multivariate models and multivariate dependence concepts*, CRC Press.
- Killiches, M., Kraus, D., and Czado, C. (2015), “Model distances for vine copulas in high dimensions with application to testing the simplifying assumption,” *arXiv preprint arXiv:1510.03671*.
- (2016), “Examination and visualization of the simplifying assumption for vine copulas in three dimensions,” *arXiv preprint arXiv:1602.05795*.
- Koenker, R. (2005), *Quantile regression*, no. 38, Cambridge university press.
- (2013), *quantreg: Quantile Regression*, r package version 5.05.
- Koenker, R. and Bassett Jr, G. (1978), “Regression quantiles,” *Econometrica: journal of the Econometric Society*, 33–50.
- Koenker, R. et al. (2011), “Additive models for quantile regression: model selection and confidence band-aids,” *Brazilian Journal of Probability and Statistics*, 25, 239–262.
- Kotz, S. and Nadarajah, S. (2004), *Multivariate t-distributions and their applications*, Cambridge University Press.
- Li, Q., Lin, J., and Racine, J. S. (2013), “Optimal bandwidth selection for nonparametric conditional distribution and quantile functions,” *Journal of Business & Economic Statistics*, 31, 57–65.
- Meinshausen, N. (2006), “Quantile regression forests,” *The Journal of Machine Learning Research*, 7, 983–999.
- Nagler, T. (2015), *kdecopula: Kernel Smoothing for Bivariate Copula Densities*, r package version 0.0.1.

- Nagler, T. and Czado, C. (2015), “Evading the curse of dimensionality in multivariate kernel density estimation with simplified vines,” *arXiv preprint arXiv:1503.03305*.
- Nelsen, R. B. (2007), *An introduction to copulas*, Springer Science & Business Media.
- Noh, H., Ghouch, A. E., and Bouezmarni, T. (2013), “Copula-based regression estimation and inference,” *Journal of the American Statistical Association*, 108, 676–688.
- Noh, H., Ghouch, A. E., and Van Keilegom, I. (2015), “Semiparametric Conditional Quantile Estimation through Copula-Based Multivariate Models,” *Journal of Business & Economic Statistics*, 33, 167–178.
- Parzen, E. (1962), “On estimation of a probability density function and mode,” *The annals of mathematical statistics*, 1065–1076.
- R Core Team (2015), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
- Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., and Erhardt, T. (2015), *VineCopula: Statistical Inference of Vine Copulas*, r package version 1.7.
- Sklar, M. (1959), *Fonctions de répartition à n dimensions et leurs marges*, Université Paris 8.
- Spokoiny, V., Wang, W., and Härdle, W. K. (2013), “Local quantile regression,” *Journal of Statistical Planning and Inference*, 143, 1109–1129.
- Stöber, J., Joe, H., and Czado, C. (2013), “Simplified pair copula constructions – limitations and extensions,” *Journal of Multivariate Analysis*, 119, 101–118.
- Wu, T. Z., Yu, K., and Yu, Y. (2010), “Single-index quantile regression,” *Journal of Multivariate Analysis*, 101, 1607–1621.
- Xiao, Z. and Koenker, R. (2012), “Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models,” *Journal of the American Statistical Association*.